Waves in space

In an open circuit, only an electric field is detectable. Is this because there is no magnetic field present?

by Ivor Catt

Prior to Maxwell, a great deal of theory had been developed around electric and magnetic fields. This theory included Kirchhoff's Laws, the Biot-Savart Law and Ampère's Rule. Electrical circuits were generally steady state, or at worst slowly varying, and the problem of whether electrical and magnetic effects traversed distance instantaneously or took time to propagate did not arise.

Because fields were steady or slowly varying, experiments were generally limited to the study of closed circuits of conductors (and resistors). However, capacitors (electrolytics) were also used, and these created an anomaly in a theoretical structure which included Ampère's circuital law ($\oint H dl = i$) and Kirchhoff's second law ($\Sigma i = 0$). When the switches were closed, electric current flowed in the loop and (following Ampère's circuital law, also called Ampère's Rule,) magnetic flux appeared in the space around the wires.

Ampère's Rule says that if we describe a closed loop, the line integral of the magnetic field strength along the edge of the loop is related to the electric current through any surface bounded by the loop.

The capacitor created an anomaly, because a closed loop could be described where i had more than one value, depending on whether the surface (S_1) cut the conductor or S_2 passed between the plates of the capacitor. Consequently the absurd situation arose that $\oint H dl$ had to have two values at the same time.

Maxwell 'cut the Gordian Knot' by asserting that the rate of change of electric field between the capacitor plates behaved just like a real current i. So Ampère's rule became

$$\oint H dl = i + \int_S \frac{dD}{dt} ds$$

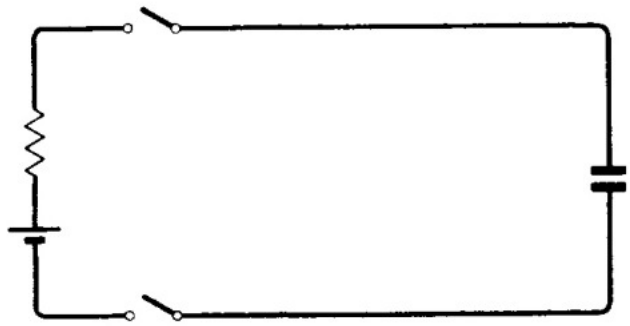
It is important to remember that the *premise* which preceded the problem of the capacitor was that electric currents and fields were steady or slowly varying. It was accepted that, at the moment the switches closed, the current i appeared at all points in the circuit. The time for the effect of the switch closure to travel across the distance from switches to capacitor was zero.

WIRELESS WORLD MARCH 1983

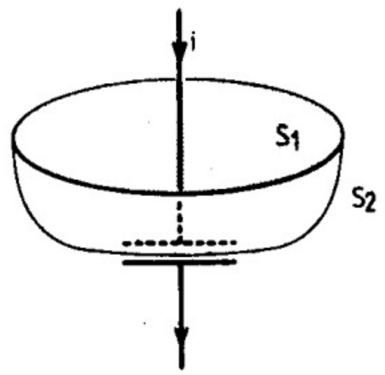
The current-like field dD/dt between the capacitor plates led Maxwell to conjecture that there could be electromagnetic 'waves in space'. It was already known that a changing magnetic field produced electric current (Faraday's Law $v=d\phi/dt$) and that electric current produced a magnetic field (Biot-Savart Law $H= i\, dl \sin\theta/4\pi r^2$.) The changing electric field dD/dt seemed to be an electric current in space. With both changing magnetic fields and changing electric currents in space, we seemed to have the possibility of wire-less propagation of electromagnetic signals using a crabwise progression of cause and effect; electric current → magnetic field → electric current

The error in this whole business occurred right at the start. Let us assume that the conductors linking battery to capacitor are one light year long. When the switches are closed, it is obvious that current will not immediately flow in the capacitor. A wave front must travel from switches to capacitor, and behind that front will be electric field and magnetic field – we have a transmission line. Also, should the distance between the two conductors or their shape change, some of the wave front will continue to the right and some will reflect to the left, carrying back the message about the change.

The front end of the capacitor is merely one such change in the cross section of the transmission line. The far (open circuit) end of the capacitor is another such change.



(a)



(b)

Fig 1. *The elemental closed loop (a), in which the capacitor creates the difficulty that the current must have two values, depending on whether the surface (b) cuts the conductor or bisects the capacitor.*

The problem Maxwell should have been concerned about was how the electromagnetic field developed between the wires when the switches were closed, not what happened in a capacitor. The transmission line problem (AB) *precedes* the capacitor problem (CD), and the capacitor problem would be solved automatically with the solution of the transmission line problem.

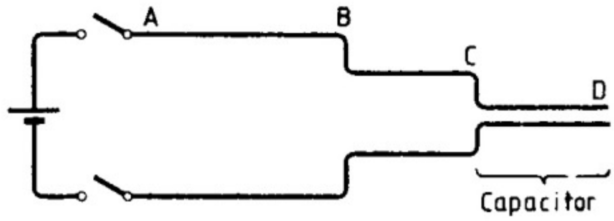


Fig 2. *A changing separation between conductors reflects energy. The capacitor is simply another change.*

Before the switches are closed, we can measure a voltage and electric field but find no trace of a magnetic field. When the switches are closed, an electric current starts off down the wires and a magnetic field begins to appear between the wires. The conclusion that the voltage (or pressure) causes the electric current which in turn causes the magnetic field is compelling, and it is not surprising that this mistaken view has lasted for a century. It is then a short step to say that the changing magnetic field in its turn (by Faraday's Law) generates an electric field and thence a (displacement) current, and the sequence can start again.

But was there *really* no magnetic field before the switches were closed?

Let us consider a steady charged capacitor. Does it have no magnetic field, only an electric field? In order to understand the situation in the battery and wires up to the switches before they are closed, it is useful to study the reed-relay pulse generator.

The reed relay pulse generator was a means of generating a fast pulse using rather primitive methods. A one metre section of 50 ohm coax. AB was charged up to a steady 10 volts (say) via a one megohm resistor, and then suddenly discharged into a long piece of coax. BC by the closure of two switches.

A five volt pulse two metres wide was found to travel off to the right at the speed of light for the dielectric on closure of the switches, leaving the section AB completely discharged.

(The practical device lacked the second, lower switch at B, which is added in the diagram below to simplify the argument.)

The curious point is that the width of the pulse travelling off down BC is twice as much as the time delay for a signal between A and B. Also, the voltage is half of what one would expect.

It appears that after the switch was closed, some electromagnetic energy must have started off to the *left, away* from the now closed switch; bounced off the open circuit at A, and then returned all the way back to the switch B and beyond.

This paradox, that when the switches are closed, electromagnetic energy promptly rushes *away* from the path suddenly made available, is understandable if one postulates that a steady charged capacitor AB is not steady at all; it contains electromagnetic energy, half of it travelling to the right at the speed of light, and the other half travelling to the left at the speed of light.

Now it becomes obvious that when the switches are closed, the rightwards travelling electromagnetic energy will exit down BC first, immediately followed by the leftwards travelling electromagnetic energy after it has bounced off the open circuit at A. Even before the switches were closed, every segment of electric field had coexisting with it a segment of magnetic field at right angles, and both were travelling together at the speed of light.

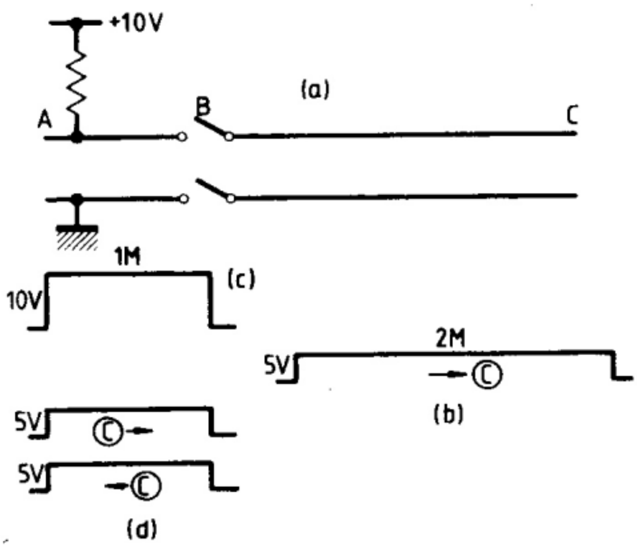


Fig 3. *Pulse generator (a). Changing AB to 10V and closing switches produces 5V pulse twice as long as AB (c) and (d), caused by left and right travelling pulses following down line BC.*

What is true of a 'steady' charged capacitor or coax. cable is also true of a pair of wires connected to the battery. Before closure of the switches, electromagnetic (not electric) energy was oscillating to and fro between battery and switches. Since the same amount travelled in both directions, the magnetic fields being equal and opposite cancelled, and only an electric field could be detected. 'Waves in space' existed between these two wires long before the switches were closed and before the capacitor came into the picture.

DEATH OF ELECTRIC CURRENT

I have progress to report.

D. W. Bell, who is not given to wasting words, said in his letter (October 1982) that the role of mathematics in physics "is essentially *predictive*" and concluded his letter "But if one accepts the logic of mathematics, one can accept the logic of mathematical models." It is clear from the introduction to his paper that Hertz would have agreed with Professor Bell; in fact Bell has explained the motive for every experiment performed by Hertz between 1886 and the time of his untimely death on the first day of 1894 at the age of 36. By accepting the logic of Maxwell's mathematical model of an ether, Heaviside and Poynting were the first scientists to realise that Maxwell's equations predict that the source of a current in a wire was located in the surrounding field. Hertz agreed with the mathematical reasoning of the Heaviside-Poynting theory "as the correct interpretation of Maxwell's equations."

Catt's critics, although not accepting the logic of Maxwell's mathematical model, have all based their criticism on the fact that Maxwell's equations predict the phantom existence of his displacement current. Maxwell's own definition of his displacement current is in Art. 111 of his Treatise, dealing with the phenomenon of induction of electricity through non-conductors.

> "*Electric Displacement*. When induction is transmitted through a dielectric, there is in the first place a displacement of electricity in the direction of the induction. For instance, in a Leyden jar, of which the inner coating is charged positively, and the outer coating negatively, the direction of the displacement of positive electricity in the substance of the glass is from within outwards.
>
> Any increase of this displacement is equivalent, during the time of increase, to a current of positive electricity from within outwards, and any diminution of the displacement is equivalent to a current in the opposite direction."

In other words, only during an acceleration or deceleration of the velocity of electric displacement does Maxwell's displacement current manifest itself. Maxwell said in Art. 62 that all electric currents flow in closed circuits, and in Art. 305 that as all currents of conduction must flow from a high to a low potential, conduction currents cannot flow in closed loops. I have suspected that all current loops are closed, and more importantly caused by, a displacement current, for instance in the induction of electricity from the primary to the secondary winding of a transformer. Hertz's paper seems to confirm this is so. The present confusion in electromagnetic theory lies in our failure to differentiate between electric displacement and displacement current; the latter only manifests itself when the momentum of the former either accelerates or decelerates.

Ivor Catt's Heaviside Signal or Poynting Vector travels through space at the constant velocity of light, and is therefore by Newton's first law of motion, inert. It is a form of perpetual motion, and will travel through space at its constant velocity forever, unless acted upon by a polarized force. Newton defined inertia as a 'latent' or potential force. If a body at rest or travelling at a constant velocity is either accelerated or decelerated, its equal and opposite reaction to a polarized force causes its latent force to be transformed into an active force, because a force is the product of a mass and an acceleration or deceleration. Maxwell's electric displacement also travels through his ether at the constant velocity of light in free space in the form of a wave of displacement or strain of his ether, and like the Heaviside Signal, will do so forever unless a polarized force, such as a conductor, decelerates the electric displacement and changes it into a displacement current. When the displacement of the potential energy of the ether is accelerated from a state of rest to the velocity of light, the resultant strain is in the form of a displacement current during the period of accelera-

tion. When a wave of electric displacement of the intensity of the ether's potential energy suffers a deceleration after its flight through space at a constant velocity, the electric displacement's kinetic energy is transformed into an electromotive force which produces a displacement current. The e.m.f. causes a displacement current to penetrate the surface of a conductor of electricity, say an aerial.

In the case of very-low-temperature superconductivity, I believe Maxwell's equations and his mathematical model predict that the wire presents an impenetrable barrier and perfectly frictionless surface of slip to the electric displacement in the neighbourhood of the wire, and the current is inert and flowing in a closed loop at a constant velocity in the surrounding field only. As the temperature of the wire increases, the wire's surface loses its properties, and the reactive centripetal force of the surrounding ether aimed at the centre of the wire, decelerates the momentum of the electric displacement by forcing it to penetrate the surface of the wire, producing a displacement current in the wire. The permittivity, or modulus of electric elasticity of the ether surrounding the individual atoms of the mass of the wire must decrease as the wire's temperature increases. The flow of heat is a form of displacement current.

Hertz's paper raises many questions which are sure candidates for the immediate application of Dr Murray's Doctrine of the Improper Question. If a current of conduction is caused by the penetration into the wire by displacement current, is the current when steady, travelling at a constant velocity longitudanally through the length of the wire, or, as Maxwell's equations predict, acting vertically through the surface of the wire only?

Should we call the electric current in a conductor the Catt Effect?

M. G. Wellard
Kenley,
Surrey

I refer to the letter from Mr Ivor Catt in the *WW* for February 1983. He asked me to look at his diagram on p.80 *WW* December 1980. I have now been able to do this, courtesy of the *WW* reprint service.

It has taken me several days (and sleepless nights) to see what was in his mind, and do not mind admitting I got off to what I think was a false start in what I intended to say by reply, because I think he has made a mistake in what he invites me to do. So if he does not mind I am going to do two things my way.

Firstly, that 50ohm bit that he wants to put in the upper plate; I am going to do so loosely, so that it can be removed without touching it, by means of a sudden surge of gravity, or a puff of wind, or an angel on wings, so that whatever portion of the total charge is residing on it goes with it, leaving a gap in the surface. What was one charged capacitor is now two smaller ones, each carrying less than half the original charge.

Secondly I am not, in the interests of simplicity, going to use a length of coax., but rather to employ two parallel conductors of a spacing which entitles them to the nominal qualification of 50ohms, erected in the way he asks for. What have I got now? No more or less than two terminal posts, one for each capacitor, each of the same sign and potential.

We can do as we please in the way of rearranging these charges from external sources.

What we have not got is a pair of conductors so placed and utilized that they can be said to be exhibiting a Z of 50 ohms to any external influence. So they are not by my reckoning an accurate substitute for the 50ohm resistor we got the angels to take away.

What I will join in and say, is that of course in charging and discharging these two capacitors, or the original one for that matter, at the velocity of light or thereabouts we do have a time lapse from terminal to the most remote part of the conducting surfaces concerned, which does not help me to consider the behaviour of frictionally induced charges on insulators.

O. Dogg
Hurstpierpoint,
Hussocks,
West Sussex.

WIRELESS WORLD MAY 1983

HERETICAL PHYSICS

Those of us who are approaching the age of 80 can hardly bear to wait a month to find out what kind of Newer Physics is going to turn up in the *WW* "Letters". Some of the ideas are so oddly fascinating that it seems a real pity that they cannot all be right.

Attractive though this is, I cannot help remembering the professor of physics who reminded his students that what happened on the lab. bench was real, whereas what went on inside human heads was mostly fantasy, and often pathological fantasy at that.

P. C. Smethurst,
Bishop's Stortford, Herts.

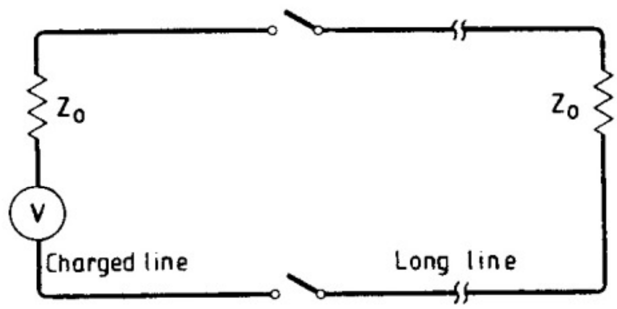
WAVES IN SPACE

Ivor Catt (March, 1983) says "the voltage is half of what one would expect". The curious point to me is on what he bases his expectations. If the charged line is regarded as a voltage source of impedance Z_0 (the characteristic impedance) connected to another impedance Z_0 through the switches, as in the accompanying sketch, then the voltage is exactly what one would expect. The finite duration of the pulse stems from the fact that the charged line is an energy storage device (electrostatic field) and not a source of e.m.f. which implies energy conversion. It is worth adding that the impedance technique also enables one to predict occurrences when the Z_0 of the long line is not equal to that of the charged line.

As to the claimed paradox that "electromagnetic energy promptly rushes away from the path suddenly made available", if Mr Catt will examine Poynting's vector in the charged line after the switches are closed, he will find that the electromagnetic energy moves only into the long line, that is, from left to right.

It is possible to regard the condition of the charged line as the result of interference between two waves travelling in opposite directions, just as one can treat a straight line as the arc of a circle of infinite radius. There are times when they are useful models of reality. The real paradox of the article is the question of where Mr Catt (and, for that matter, *Wireless World*) have been all this time. To my knowledge, the Royal Air Force used this approach to transmission lines as pulse generators in 1959, and I have no doubt that the technique goes back much further in time.

R. T. Lamb
British Telecom
Milton Keynes



WIRELESS WORLD AUGUST 1983

I refer to Mr Catt's article in the March 1983 issue.

In a letter, I could not hope to reproduce the great body of scientific and engineering knowledge that has amply demonstrated the non-relativistic interpretation of Maxwell's Equations or of Einstein's treatment using special relativity. If the theory is so seriously flawed it is surprising that we can design and build antennas and microwave devices. Nevertheless, I cannot let Mr Catt's analysis of te pulse generator go unchallenged, especially as it is so easy to demolish his arguments.

Firstly, if a piece of charged coax. really has equal and opposite waves running in each direction why are they not attenuated by the losses in the line? After all in one second each of his waves would have travelled nearly 2,000 miles in lossy coax.

Secondly, if I connect an antenna to a piece of coax. I can still charge up the line. Why do not these waves of which he speaks radiate into space? Or, at least, the high-frequency components of the pulse to which the antenna will be matched.

The conventional solution to his 'exotic' problem can be found by solving the transmission line equations for a cable under the stated starting conditions (see reference). As this is rather tedious and since Mr Catt seems to prefer hand waving to mathematics I will at least demonstrate where the 2m pulse length comes from.

When the extra length of line is connected to the 1m line, charge starts to move down the new line charging the distributed capacitance of the line through its distributed inductance as it goes. (Maxwell's equations applied to circuits show us that capacitors connected together share their charge). This leaves a void which the 10V (1m) line fills. The void propagates towards the open end of 1m line at the speed of the line. The charge close to the open end of the line will be liberated at a time equivalent to 1m of line and will take 1m to propagate to the other end, explaining the 2m length of the pulse.

The pleasing aspect of the above argument is that we do not have to destroy a century of succesful electro-magnetic theory to produce it. If Mr Catt has so much more insight into electro-magnetic theory than the rest of us it is surprising that he has not produced any new microwave devices that demonstrate his superior understanding.

Timothy C. Webb
Columbia, MD
USA

Reference: Brown and Glazier: Signal Analysis pp. 345-349.

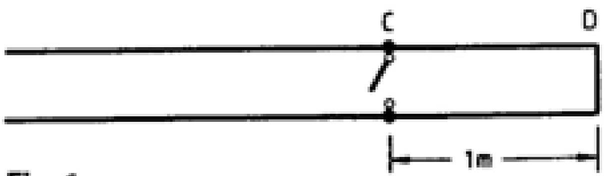


Fig. 1.

I was not too sure whether I should be amused or startled by Ivor Catt's article on "Waves in Space" in which he postulated the existence of electromagnetic energy in a static electric filled, when I reminded myself that the date was 1st April. It had never occurred to me, even remotely, that a magnetic field could be directly caused to exist by the presence of an electric field. However, being sympathetic to the idea that all things appear to be possible in this day and age, I allowed my mind to be bent a little further and read on.

The production of a voltage pulse in a transmission line which is half the amplitude and twice the length of that existing statically in piece of coaxial cable, can form the basis of a number of interesting experiments. For example, Mr Catt's travelling pulse can be converted back into a static charge again if his coaxial cable is terminated into an open circuit but with a pair of switches 1 metre from the end as shown in Fig. 1. When the switch at B is closed and the resulting pulse eventually reaches D, it will of course, be reflected (double back on itself); however, if switch C, is opened at the instant the leading edge meets the trailing edge, no great drama ensues but we are left with a 1 metre piece of statically charged coaxial cable as before. It is also interesting to consider what happens if point D is terminated into a short circuit. This time, the pulse will be converted into one of twice the current at zero voltage and the leading and trailing edges will be locked together and oscillate back and forth converting the pulse between a current at zero voltage and a voltage at zero current until this activity decays due to losses.

However, perhaps it would be more interesting to consider what would happen if this final 1 metre of coaxial cable is made superconducting and instead of isolating switches at C, a short circuiting switch is provided as shown in Fig. 2. If the switch at C is closed at the instant the leading and trailing edges meet, electrons at zero voltage, will continue to flow around the 1 metre coaxial circuit as a direct current. It would, perhaps, be better to say that current drifts around, because depending on the construction of the coax., it could take hours for any single electron to work its way around the circuit. It should, of course, be remembered,

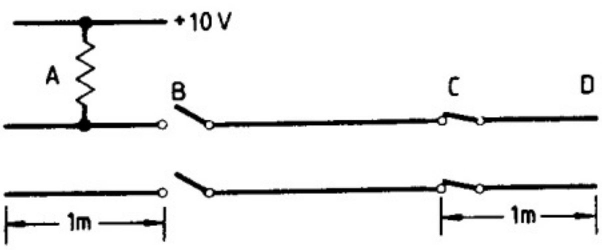


Fig. 2.

that as no voltage is present, the density of the electrons in the closed circuit is the same as that in the remainder of the uncharged transmission line.

Unless someone comes along and throws into doubt that an electric field exists between adjacent electrons and protons it is easy to demonstrate that in the above circuit condition, countless billions of electric fields exist alongside the magnetic field caused by the current, although no measurable voltage exists. Similarly in the case of the purely electrostatic condition, the same large number of magnetic fields exist due to the rotating and spinning electrons. However, I doubt somehow whether this has anything to do with the way the charged section of coaxial cable propagates a pulse whose length is twice that of the charged section. And I must confess, that the prospect of opposing magnetic fields oscillating back and forth along the cable as Mr Catt suggests, appears to be even less likely.

Surely, the answer to his paradox is simply that on closing the switch at B, the electrons that flow out of the negatively charged conductor and those that flow into the positive conductor of the cable not only cause a wave front of current and voltages to be transmitted along the line but also back into the charged section itself. And since the energy has to be shared between both fronts, the voltage will be halved. The discharging current flowing into the charged section will set up a magnetic field, which, on collapsing, will produce an equal pulse of voltage and current to follow on the heels of the pulse which has already departed from he originally charged section.

R. J. Hodges
Bath
Avon

DEATH OF ELECTRIC CURRENT

I believe Ivor Catt bases his theory on Heaviside's "the current in the wire is set up by the energy transmitted through the medium around it."

Chapter ten of Hertz's book 'Electric Waves' is a reprint of his paper 'On the Propagation of Electric Waves by Means of Wires' first published in 1889, a year after the experiments which made him famous. The purpose behind the experiments described in this later paper was to test Heaviside's and Poynting's theory that, as Hertz wrote, "the electric force which determines the current is not propagated in the wire itself, but under all circumstances penetrates from without into the wire. . . ." Hertz went on to say "As a matter of fact the theory was found to be confirmed by the experiments which are now to be described; and it will be seen that these few experiments are amply sufficient to support the conception introduced by Messrs Heaviside and Poynting."

Hertz then described a set of experiments which used his invention of the coaxial cable and the balanced feeder or transmission line, and concluded his paper, "On studying the experiments above described, the mode in which we have interpreted them, and the explanations of the investigators referred to in the introduction, one difference will be found especially striking between the conception here advocated and the usually accepted view. (Weber's theory of electricity carried by charged particles acting instantaneously at a distance.) In the latter, conductors appear as the only bodies which take part in the propagation of electrical disturbances – non-conductors as bodies which oppose this propagation. According to our conception, on the other hand, all propagation of electrical disturbances takes place through non-conductors; and conductors oppose this propagation with a resistance which, in the case of rapid alterna-

tions, is insuperable. We might almost feel inclined to agree to the statement that conductors and non-conductors should, according to this conception, have their names interchanged . . . "

Hertz was even more specific in his Supplementary Note No. 24. "By the experiments in the following paper it is pretty plainly proved that in the case of rapid variations of current the changes penetrate from without into the wire. It is thereby made probable that in the case of a steady current as well, the disturbance in the wire itself is not, as has hitherto been assumed, the cause of the phenomena in its neighbourhood; but that, on the contrary, the disturbances in the neighbourhood of the wire are the cause of the phenomena inside it."

Catt's critics have a choice: either Hertz was a crank and a crackpot, or he was, as an experimenter and detective, in the same class as Faraday. If Hertz's diagnosis of his experiments with a transmission line is correct, the effect we call a current is caused by "the disturbances in the neighbourhood of the wire," what, in the neighbourhood of the wire, is being disturbed? Maxwell's ether?

M. G. Wellard
Kenley Surrey

WAVES IN SPACE

I refer to the correspondence in the August issue concerning Catt's "Waves in Space", (March, 1983).

Ivor Catt has, for some years, been proposing new explanations of electrical phenomena which many regard as already fully explained by classical e-m theory; theory which unfortunately has become dogma because few have bothered to question its tenets in those areas where its teachings give rise to curious and unexplained paradoxes.

Correspondents who try to put Catt down generally throw up a dogmatic smokescreen whilst often failing competely to address themselves to the apparently paradoxical events he has attempted to explain. The latest correspondence concerning "Waves in Space" is no exception. R. T. Lamb's letter gives no real explanation of the phenomena that Catt discussed and simply fluffs the issue of pulse duration with a remark about the charged line being an energy storage device rather than a source of e.m.f.

Timothy C. Webb's letter puts a finger on one important issue when he asks why Catt's contra-moving waves are not destroyed by line losses, but he fails to ask whether conventional energy dissipation due to line loss applies to these contra-moving waves. I suspect Catt thinks otherwise and it would be interesting to have his views.

In other respects Mr Webb's letter falls into the dogma trap. There is a resounding bit about "The great body of scientific and engineering knowledge that has amply demonstrated . . ." etc., etc. Dr Catt has quite reasonably asserted that the great body of scientific and engineering knowledge has singularly failed amply to demonstrate some of the things it purports to explain! At the end of his letter Webb gives what I found to be an incomprehenisble explanation of the pulse duration problem and then rounds this off with a remark about the "pleasing aspect of this argument . . ." etc. I was not very pleased because I could not make head nor tail of it!

Hodge's letter is perhaps more thoughtful but again it does not seem to explain the phenomena which Catt discussed in his article.

Catt's theories may be wrong but he is certainly right to shine lights into some of the dark and deceiving corners of classical e.m. theory. One would like to see more reasoned arguments advanced in refutation and less reliance on the "dogma must be right" approach which, incidentally, rather neatly mirrors the discussion on the "closed loop arguments", (said to be used to support relativistic dogma), given in an unrelated letter from A. H. Winterflood in the same issue of *Wireless World*.

Anyone who thinks he knows all about electricity should also read Professor Jennison's article on making a charge from a radio wave!

M. G. T. Hewlett
Midhurst
W. Sussex

DEATH OF ELECTRIC CURRENT

Two quotations from *Wireless World*, September 1983:

1. M. G. Wellard's letter, quoting Hertz: "In [Weber's theory], conductors appear as the only bodies which take part in the propagation of electrical disturbances – non-conductors as bodies which oppose this propagation. According to our [Hertz's] conception . . . all propagation of electrical disturbances takes place through non-conductors; and conductors oppose this".

2. M. McLoughlin: "Current dumping review – 1": "Obviously high farce has effected an entry . . . A complex situation may sometimes be viewed quite validly in alternative ways. In this case the fullest understanding seems to be obtained when one has seen both explanations, seen that they are both valid, and grasped that they are complementary views of the same situation".

Need one say more?

R. Kennaway
Norwich

Advance into the past

Ivor Catt looks at the inhibitions imposed on designers of computers by the conventional mythology of devices and architecture.

by Ivor Catt

Rational forward progress in computer technology could only be achieved if a significant proportion of computer scientists had some mastery of most of the technologies and disciplines involved. Unfortunately this is not the case, because the necessary spread of knowledge and understanding – from semiconductor physics at one extreme to complex software and computer applications at the other – is too broad.

Computer scientists habitually assume that the conventional wisdom, or myth, imposed on other specialities than their own, is true. They find it convenient to base their views on the state of the art in other fields on information supplied by amateurs rather than those actually working in them. A specialist in any one field tends to see his professional survival as depending on the stabilization of the conventional-wisdom straight-jacket which at one time or another has been imposed on every other speciality. This is because change in these other fields would make his own speciality too fluid, and he would not survive . . . a point of view which, although usually subconscious, sometimes comes out into the open.

For example, around 1970 it was commonly said, "We are having so much difficulty mastering the software of present computers that it is important, *if we are to progress*, that computer hardware be frozen for a decade or more." Some readers will see the irony implicit in this comment, which was often made by programmers with no knowledge of engineering, which meant virtually all programmers. There followed an explosion of complex software techniques, including list processing, which could have been much more easily achieved by hardware modification; but this option had been outlawed. The result was an increase in the complexity and confusion of already over-complex software, and a deterioration in the overall position.

In general, all other disciplines ganged up on each individual discipline and forced it to remain essentially static, at least in its perceived structure when it interfaced with other disciplines. Examples are:

- The blocking of any blending of memory and processing, any move away from absolute von Neumann, and strict adherence to the 'von Neumann bottleneck', even though at one extreme the technology was demanding it and at the

other extreme almost all applications were demanding it.

● The blocking of any deviation from the traditional drift from fully serial machines to fully word-parallel machines, even though (a) the technology demanded a reversal towards serial working, (b) the change in the relative cost of circuit and interconnection demanded it, and, strongest of all, (c) a strong mythology had developed that the computer industry was combining with an avowedly serial industry, telecommunications (citing the appointment of a Minister for Information Technology as evidence). Here we see one myth combating and overcoming another, unfortunately the wrong one, "fully parallel fetishism", being the victor.

● The imposition for all time of the t.t.l. logic signal as industry standard. This occurred even though t.t.l. logic, which came into general use in spite of its weaknesses in design (including the heavy standing current in signal lines, the high signal swing, etc), had given way as the industry standard circuit to c.m.o.s., which had much greater circuit density, in which a very different logic signal standard would have been more efficient.

● The maintenance of a key feature of the thermionic valve – the idea that hermetic seal was necessary to stop the cathode from burning up – well beyond the disappearance of the cathode through drastic changes in the technology towards silicon semiconductor l.s.i. Few computer engineers realise that the 'hermetic seal fetishism' which continues today in v.l.s.i. chips dates back to the danger of allowing oxygen to reach a hot cathode, and has nothing to do with semiconductor technology.

● The inexplicable standardization, without a murmur, on the use of Kovar as the metal for the leads coming from an integrated circuit chip, even though every parameter of Kovar except one is bad in this application. The one good parameter is that Kovar wets to glass, so allowing the formation of a hermetic seal. Kovar's bad features include the following:

– It has rather high electrical resistivity, so degrading performance by creating extra voltage drop in the signals entering a chip.

– It is magnetic, so that signals into a chip are delayed, and energy wasted, while the magnetic field is built up.

– It is not ductile, and work hardens fast, so that there is an unnecessarily large risk of fracture due to bending or vibration.

– Worst of all, it does not wet to solder. In order to make it possible to solder to a Kovar lead, the lead has to be gold plated. However, during the soldering process, the gold dissolves into the solder, creating a brittle alloy and also, should soldering and de-soldering be repeated, the dissolving away of all the gold and the creation of a dry, non-wetted joint between solder and virgin Kovar.

● Microprocessor manufacturers have displayed ignorance of the mechanism of digital signal propagation and voltage decoupling. Placing voltage pins at opposite corners of the package, thus introducing a large single-turn inductor in series with the voltage supply, is the worst possible pin choice, limiting the speed of microprocessors and also making them pattern-sensitive. Although only marginally significant in the old 14 or 16-pin dil integrated circuit, the problem created increases rapidly as the square of the package length, mak-

ing the large microprocessor chip slow (only 4 MHz), pattern sensitive, and depdendent on the layout of the host printed-circuit board in a manner not understood (and so not predicted) by system designers.

● Looking at another aspect of the standard package, I suppose that I should be relieved that the industry did not standardize on the even more absurd IBM SLT package, 1965 vintage, which had a line of pins down all four sides of a square package. When deciding how the pins should exit from an integrated circuit package, the decisive aim should be to minimize the obstruction of printed circuit conductors in the host p.c.b. The two, unrelieved, lines of pins are about as obstructive, and therefore as inefficient, as it is possible to devise (*pace* the IBM SLT). Alternate pins should have been staggered, and this is a simple operation (which would not have created significant problems in the manufacture of i.c. sockets). I only mention this to show how thoughtless and casual developments have been, not to propose change at this late stage.

● The standardization by the industry of t.t.l. with its totem pole (push-pull) output was based on the mistaken idea that the load seen by an i.c. output is capacitive. This was true for thermionic-valve logic gates, with their high impedance, low current outputs, but ceased to be true when we used transistors, at which point the load seen by a fast output became resistive; either a transmission line characteristic impedance (resistance) or a t.t.l. input load (also essentially resistive). Whereas a capacitive load could helpfully be driven push-pull, today a resistive load can perfectly well be driven by one transistor, as is demonstrated by the fact that the fastest existing circuit, 1 ns e.c.l. has a single transistor output.

Speed of logic

Generally, the limiting factor in the speed of logic is not the time taken for a transistor to switch on or off, but rather the time taken thereafter for the switched current to charge or discharge the stray capacitance in the line connecting this transistor to the next. A good measure of the delay involved, i.e. the gate delay, is gained by multiplying the resistance of the drive transistor when switched on by the stray capacitance that it has to drive.

When a bipolar transistor, as used in a t.t.l. circuit, is switched on, its resistance is less than 10 ohms. The capacitance of the line, or wire, on the printed circuit board joining this output to the next logic element is of the order of 20 picofarads. Multiplying these two together gives us a time delay of 200 picoseconds. This shows us that, from this point of view at least, sub-nanosecond logic speeds are possible and we do not pay a speed penalty if our logic signals skip from chip to p.c.b. to chip to p.c.b. and so on.

In stark contrast, the smallest possible unipolar, or mos transistor, when switched on, still has a resistance of 10,000 ohms. If it drives 20 picofarads of capacitance on a printed circuit board, the delay, or signal rise time, resulting would be 20 pico multiplied by 10,000, that is, 200 nanoseconds. So if the physically smallest possible (i.e. square) cmos output transistor has to drive a signal off the chip onto the printed-circuit board, the achievable speed is only 200 nanoseconds, that is, one thousand times slower than bipolar t.t.l. This dire situation can be improved by making the drive transistor bigger and so reducing its resistance. Actually, we might put ten square transistors in parallel to reduce the resistance from 10,000 ohms to

1,000 ohms. However, the price we pay is that these drive transistors have to be made very big, consuming large areas on the surface of the silicon chip. This undermines the reason for using mos which is that an mos circuit takes up less area on the chip than does a bipolar. By the way, if we make the output transistor more beefy, we can make the mos output t.t.l. compatible, and this is usually done.

Let us now consider the situation when a cmos signal on an l.s.i. chip goes from one logic stage to the next without leaving the chip. In this case, the stray capacitance which must be driven is only one tenth of a picofarad, and if the drive transistor is the smallest possible, i.e. 10,000 ohms resistance, the time constant, or delay, is only 1 nanosecond. From this we can deduce that it is not true that cmos is slow. Cmos signals across the chip have a high intrinsic speed, and so inter-chip circuitry should be serial, since this will reduce the amount of circuitry required for each function. (It is ridiculous for operations inside current microprocessor chips to be fully parallel. However, if someone made a serially operating microprocessor chip, probably nobody would buy it because, although its performance might be the same as its parallel competitors, the news would get out that the serial microprocessor contained very little hardware; there would be nothing for the salesmen to boast about.)

Note that if, by increasing the size of an output transistor by putting a number of square transistors in parallel, the output resistance of one bit of a 16-bit bus leaving the chip is brought down to 1000 ohms, so that the speed (rise time) is reduced to 20 nanoseconds, but sixteen such large transistors are needed to handle the sixteen-bit parallel word, using up valuable area on the integrated circuit surface. The same output data rate could be achieved by combining all 160 transistors in parallel to drive the sixteen bits serially down only one wire leaving the chip. In this case, assuming the same amount of chip area for the single drive transistor, a resistance of one sixteenth of 1,000 ohms could be achieved, leading to a bit rate of nearly 1,000 megabits down the single line. The point being made here is that *parallel working does not enhance speed if the circuits used are cmos*. On the other hand, a heavy price is paid when we go fully parallel – extra cost in wiring and extra pins in the i.c. package leading to extra failure (since the main cause of failure is the interconnections) and also far more failure due to pattern sensitivity with parallel data busses. Also, parallel working increases the pysical size of the resulting system, because size is largely dictated by number of interconnecting wires. It also forces us to use extremely complex, expensive test and debugging equipment including logic analysers with their awkward, octopus-like probe pods. By comparison, it is trivially easy to attach a single oscilloscope probe to a point where serial data is passing.

The Nub of computation

The heading of this section is purposedly inappropriate, to illustrate the problem at the very start. The 'computer science' discipline has come to think that its objective is 'computation', 'information processing' or some such. This is not true, or alternatively, if it is true, then 'computer science' is getting in the way of a much more important discipline, which is the application of technology to society's needs.

In our society or culture, certain historical necessities arise. It is usually thought that whether or not a certain development was a historical necessity is proven after the event by whether such a thing in fact came to pass. I think this is wrong. For instance, the wheel and axle was clearly a historical necessity in both Europe and the Americas, and the fact that the natives of the Americas never used the wheel and axle does not prove that it was not a historical necessity. More generally, we can see the extreme cases where a tribe or genus dies out because it evades a step which is a historical necessity.

Our society may well avoid historical necessity in the development of computer science, but that does not in my opinion negate the fact that what follows is a historical necessity.

The proper objective for computer science or digital electronics is to apply technology to meeting human or sociological needs. (This is a quote from my 1969 *New Scientist* article[1]') I would probably limit the broad range of application to physical, not intellectual, needs.

Any physical situation which our technology can usefully be applied to will be a multi-dimensional array of values which need (a) analysis and (b) manipulation. Digital electronics won over analogue twenty years ago, I believe for ever, and so our machinery needs to contain a *digital analogue of reality*, and in fact always does so. One measure of the elegance of our machinery, and probably of its efficiency and simplicity is the ease with which the analogue in our machine maps onto the reality of which it is an analogue. The design of an elegant (and also one suspects efficient) machine requires of the designer knowledge of the physical reality which is the target of our machine; of the nature of data manipulation and computation; and of the physical nature of the machine.

Since the ideal seems to be a machine which can be regarded as a *physical analogue of reality*, and the closeness with which the machine's structure and information mimics the physical reality, the 'computer scientist' must have competence in all fields above.

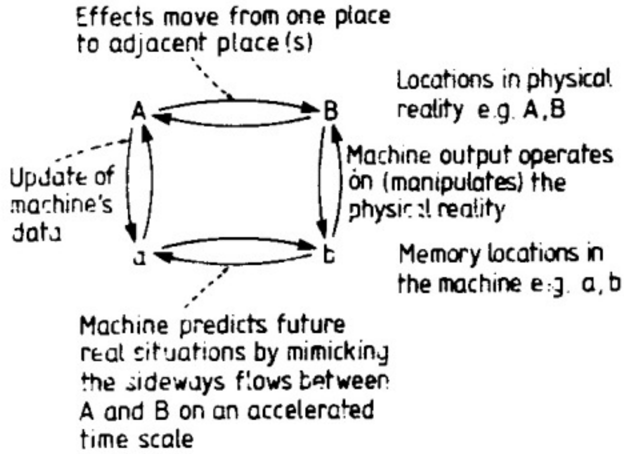
The problem is that today, programmers, calling themselves computer scientists but having no competence in anything except the second (with perhaps a little competence in the first), think they can usefully contribute to the design and development of our future machines.

A second measure of the elegance of our machinery is the degree to which changes in the physical reality we are mimicking (or recording) in our machine can be easily effected in our machine. This is why a machine is very bad if it does not have content-addressable memory, and in fact it needs more than that. It needs processing capability *in situ* in the memory. This is because values or parameters in physical reality change *in situ*, influenced only by parameters which are physically nearby. This leads to the next requirement of a good machine, which is that since in physical reality there is not action at a distance but all interaction is local, our machine should have superior (or even only) interaction capability between values (vectors, scalars, etc.) which relate to physically close points in the physical reality. Further, it appears that the ability to effect interaction between values which relate to

points which are physically distant in reality may not be necessary *at all* in our machine, although this is pushing the point rather far.

A further requirement of our machine is that updating, or interaction, capability between points in physical reality and the related points in our machine where the digital analogue for that region of physical reality is stored, should be as efficent as possible.

The task of the machine architect is to exploit the potential of his technologies to meet these requirements. I believe I have done the best compromise in the Property 1 a invention[2], but it is not ideal, considering the above criteria. The above criteria are not merely a *post hoc* rationalization



tending to show that my architectures are the best.

From the point of view of the above analysis, the reigning computer architecture theorists are doomed to failure. They (e.g. Petri nets) concentrate on the mechanism of computation in the machine, the bottom horizontal lines between a and b.

However, this has no value if the thinker does not bear in mind the dualism; that the arrows between a and b are a reflection of arrows between A and B; that computation only has value to the extent that it mimics events which occur in the real world. (This relates to my statement in the fourth paragraph above that the 'broad range of application', by which I mean the main field of application for our machine, and therefore the paradigm which should control their architecture, is directed towards practical rather than intellectual applications.

References

1. I. Catt, Dinosaur Among The Data?, *New Scientist*, 6 March, 1969

2. For Property 1 a invention, see "Wafer Scale Integration", *Wireless World*, July, 1981, pages 57-59.

WW

Page 129 is blank.

WAVES IN SPACE

I would like to make one comment on T. C. Webb's letter in the August issue. My co-author Malcolm Davidson experimented with sending data (highs and lows for 1s and 0s) in both directions down a 1 kilometre length of twisted pair. He found that the losses experienced by the signals travelling in one direction were less when pulses were being sent in the other direction.

Conventional theory would say that during the time when one positive pulse passes through another going the other way in a transmission line, the i^2R losses drop to zero. The total current is zero during this time.

Ivor Catt
St Albans
Herts

In his letter, published in *WW* November, 1983, W. M. Dalton hit a nasty land-mine that I first noticed some years ago. Let me first quote the moment when he hits it.

> "Let us start from known facts. (1) Light is an electromagnetic phenomenon: demonstrated by Faraday and Kerr. (2) Light is not a static problem: it is ocillatory (Hertz). (3) The electric and magnetic fields are at right-angles and *always* 90 degrees out of phase. Some recent textbooks show these in-phase – an unparadonable error."

I am anxious that Mr Dalton expands on why this error is unpardonable, and what disasters this error might lead us into.

First let me list some non-recent textbooks which show these in-phase.

G. W. Carter, Professor of Electrical Engineering in the University of Leeds, in his book The Electromagnetic Field in its Engineering Aspects, (Longman 1954) draws the $\bar{B}$ and E fields in-phase on page 271. Significantly, although he emphasises that E and B are at right angles (page 274) he never seems to say in the test that B and E are in phase.

A. F. Kip, Professor of Physics, University of California, Berkeley, in his book Fundamentals of Electricity and Magnetism, (McGraw-Hill 1962) draws the H and E fields in-phase on page 322. On that same page the text says that the two fields are perpendicular to each other, but does not state that they are in-phase. Again significantly, I cannot find mention in the text that they are in-phase.

O. Heaviside F.R.S., in his book Electromagnetic Theory Vol 3, 1912, in art. 452, page 4, wrote

> "The General Plane Wave . . . the slab may be of any depth and any strength, and there may be any number of slabs by side behaving in the same way, all moving along independently and unchanged. So $E=\mu vH$ expresses the general solitary wave, where, at a given moment, E may be an arbitary function of x . . "

[Replace μv by $\sqrt{\mu/\epsilon}$ – I. Catt]

Whereas some books (Carter and Kip) vaguely indicate that E and H are in-phase, other books seem to fail to discuss relative phase at all see for example Gullwick 1959, Bewley 1933. The trap was nicely set for Dalton, and he has my sympathy.

Now let us turn to my article in *Wireless World*, July 1979, entitled The Heaviside Signal.

> "We have shown that the passage of a TEM wave and all the mathmatics that has mushroomed around it does not rely on a causality relationship (or interchange) between the electric and magnetic field. Rather, they are co-existent, co-substantial, co-eternal."

In that article I compare and contrast two mutually contradictory versions of the transverse electromagnetic wave. I believe that the full realisation that E and H are in-phase deals a death-blow to one of those versions, the rolling wave, and leaves the other, the Heviside signal, the victor.

Because the differential of sin is cos and the differential of cos is minus sin, half-witted mathematicians have invaded the physics of the TEM wave and imposed a spurious story that E causes H causes E. Since sin, cos and −sin are 90 degrees out of phase, part of their phoney baggage is to imply that E and H are 90 degrees out of phase. (See my article in *WW* in March 1980.) Because the sine wave is amenable to mathematical high jinks, another part of their baggage is to imply that a TEM wave is sinusoidal. It's time we cleaned the claptrap out of electromagnetic theory.

Ivor Catt
St. Albans
Hertfordshire

TEM-WAVE PHYSICS

Lest the fierceness of Mr Catt's response to Mr Dalton (February 1984 issue) obscures what he said, could I diplomatically support all that was contained in his letter while at the same time describe a situation where E and H are 90 out of phase. This should please Mr Dalton.

But first let me remind Mr Dalton that the opposite of "static" is "dynamic" and not "oscillatory". The last is just one of many modes of motion which need not even be periodic. This is particularly important because the example I propose to give for E and H being 90° out of phase is static. This should please Mr Catt.

Starting from Maxwell's equations it is easy to derive equations of wave propagation for E and H, the solutions of which are

$$E = f(x - ct)$$

and

$$H = \frac{1}{c\mu} f(x - ct)$$

where f can be any function, not just sinusoidal or even periodic e.g. a digital (level) change, a single pulse – square or any other shape.

The variation (f) of H matches precisely the variation of E (also f) whatever whatever f happens to be. There is no delay between E and H or, in the case of f being sinusoidal, no phase difference. As Mr Catt states there is no causality between E and H. However, and this may be part of the origin of Mr Dalton's error, there is a rotation of 90° from E and H which is right handed about (not along) the direction of propagation. Thus if f is sinusoidal E and H are in phase but at right angles to each other in space, not time.

If the equations above are divided one into the other then

$$\frac{E}{H} = \frac{1}{c\mu} = \sqrt{\frac{\mu}{\epsilon}} = Z_o$$

where Z_o is the wave impedance of free space (about 375 ohms) which is independent of f.

If E and H were sinusoidal and 90° out of phase as Mr Dalton suggests, then Z_o would be the tangent i.e. from minus infinity to plus infinity. This would make it difficult for a wave to propagate. At the very least it would imply causality if one knew which occurred first and at worst would mean changing the title of your illustrious magazine.

This brings me to the example of E and H being out of phase and possibly the other half of Mr Dalton's confusion.

Suppose that a sinusoidal wave described by

$$E_1 = E_0 \sin\left\{\frac{2\pi}{\lambda}(x - ct)\right\}$$

has superimposed on it an equal wave but travelling in the opposite direction, say by reflection, described by

$$E_2 = E_0 \sin\left\{\frac{2\pi}{\lambda}(x + ct)\right\}$$

Some trigonometry reduces the sum of these to

$$E_1 + E_2 = 2E_0 \sin\frac{2\pi x}{\lambda}\cos\frac{2\pi ct}{\lambda}$$

$$\text{or } 2E_0 \sin\frac{2\pi x}{\lambda}\cos\omega$$

$$\text{Similarly} \quad H_1 + H_2 = -2H_0 \cos\frac{2\pi x}{\lambda}\sin\omega t$$

This results in the well-known standing wave where the nodes of H correspond with the peaks of E and vice versa i.e. 90° out of phase. When E is a maximum, H is zero everywhere. Then H grows and E decreases until it is a maximum and E is zero, and so on cyclically. Thus the standing wave has all the appearance of transforming itself from an entirely electric form to an entirely magnetic one and vice versa. But it is just an illusion, for as Mr Catt states, there is no causality between E and H for a single wave, still less is there any between two in which we only observe their interference pattern.

This, I hope, explains the source of Mr Dalton's confusion.

Finally I would like to disagree with Mr Catt (only in a very minor way) concerning his references. Carter in his book "The Electromagnetic Field in its Engineering Aspects" pages 266 to 276 is quite specific about there being a delay (or phase difference in the sinusoidal case) between E and H, both in his diagrams and text, and of which the above is, I trust, an accurate paraphrase. They correspond, though in different words, with the views expressed by Mr Catt.

E. O. Richards
Hitchin
Herts

PS: For those who share Mr Catt's disgust with sin and cos I commend a closer look at Walsh functions, an introduction to which appeared in these pages in January 1982. An excellent book on the subject is "Walsh Functions and the Engineering Applications".

C.A.H.